Phillippo, D. M., Dias, S., Ades, A. E., & Welton, N. J. (2020). *Assessing the performance of population adjustment methods for indirect comparisons: a simulation study*. Poster session presented at 41st Annual Conference of the International Society for Clinical Biostatistics, Krakow, Poland.

Assessing the performance of population adjustment methods for indirect comparisons: A simulation study

David Phillippo,[1] Sofia Dias,[1,2] Tony Ades,[1] Nicky Welton[1]

[1]Bristol Medical School (Population Health Sciences), University of Bristol, UK

[2]Centre for Reviews and Dissemination, University of York, UK

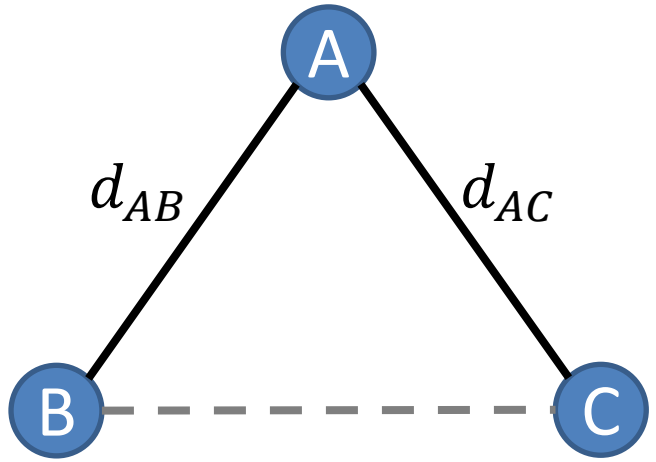
University of BRISTOL

Background – indirect comparisons and NMA

We wish to compare relative effects of multiple treatments, but often not all treatments are included in the same randomised controlled trial

Standard methods using aggregate data (AgD):

- Standard indirect comparison (Bucher IC): $d_{BC} = d_{AC} - d_{AB}$
- Network meta-analysis (NMA) extends this idea to larger networks of treatments and studies



- Assume constancy of relative effects:

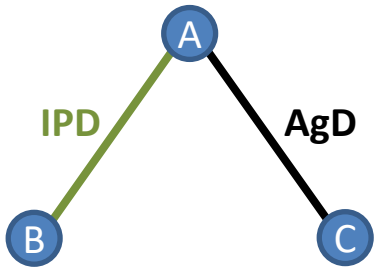
$$d_{AB(AB)} = d_{AB(AC)}$$

- Biased if there are differences in effect modifiers (EMs) between studies

Background – population adjustment

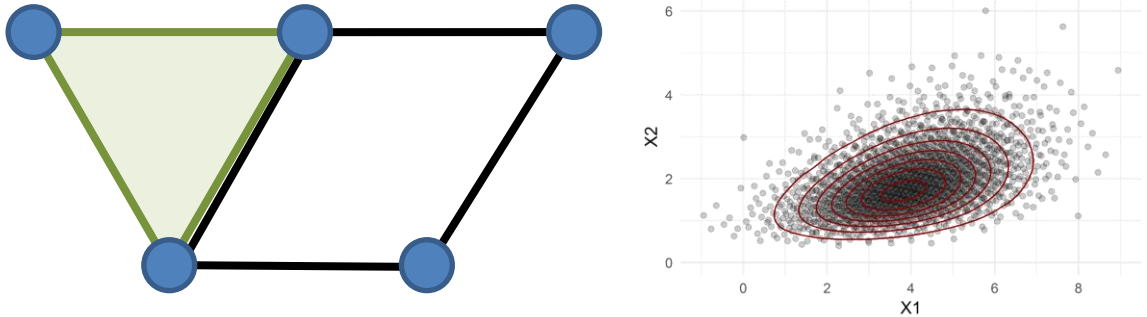
Phillippo et al. (2020)

Population adjustment methods make use of available individual patient data (IPD) to adjust for differences in effect modifiers between populations



Matching-Adjusted Indirect Comparison (MAIC) *and* Simulated Treatment Comparison (STC)

- Use weighting or regression to estimate B vs. A effect in AC population using IPD from AB study
- Limited to pairwise indirect comparisons
- Comparisons stuck in aggregate AC population



Multilevel Network Meta-Regression (ML-NMR)

- Define IPD meta-regression model ("gold standard")
- Integrate individual model over joint covariate distributions in AgD studies (numerical integration)
- Extends the standard NMA framework
- Is applicable in networks of all sizes
- Produces estimates in any target population
- Avoids aggregation bias

Simulation study

- Two study (AB – AC) setting
- Generate binary outcomes (logit model), two EM covariates

$$Y_{ijk} \sim \text{Bern}(\theta_{ijk})$$

$$\text{logit}(\theta_{ijk}) = \mu_j + q\left(\boldsymbol{X}_{ijk}^T - \boldsymbol{m}_j^T\right)\boldsymbol{\beta}_k + \gamma_k$$

$$\boldsymbol{X}_{ijk} \sim \phi_j$$

- Estimate population-average relative effects d_{AB}, d_{AC}, d_{BC} in each study population
  - MAIC, STC, ML-NMR – each with all EMs and missing an EM
  - Plus standard "unadjusted" IC (Bucher)
  - Assess bias, model and empirical SE (and coverage)

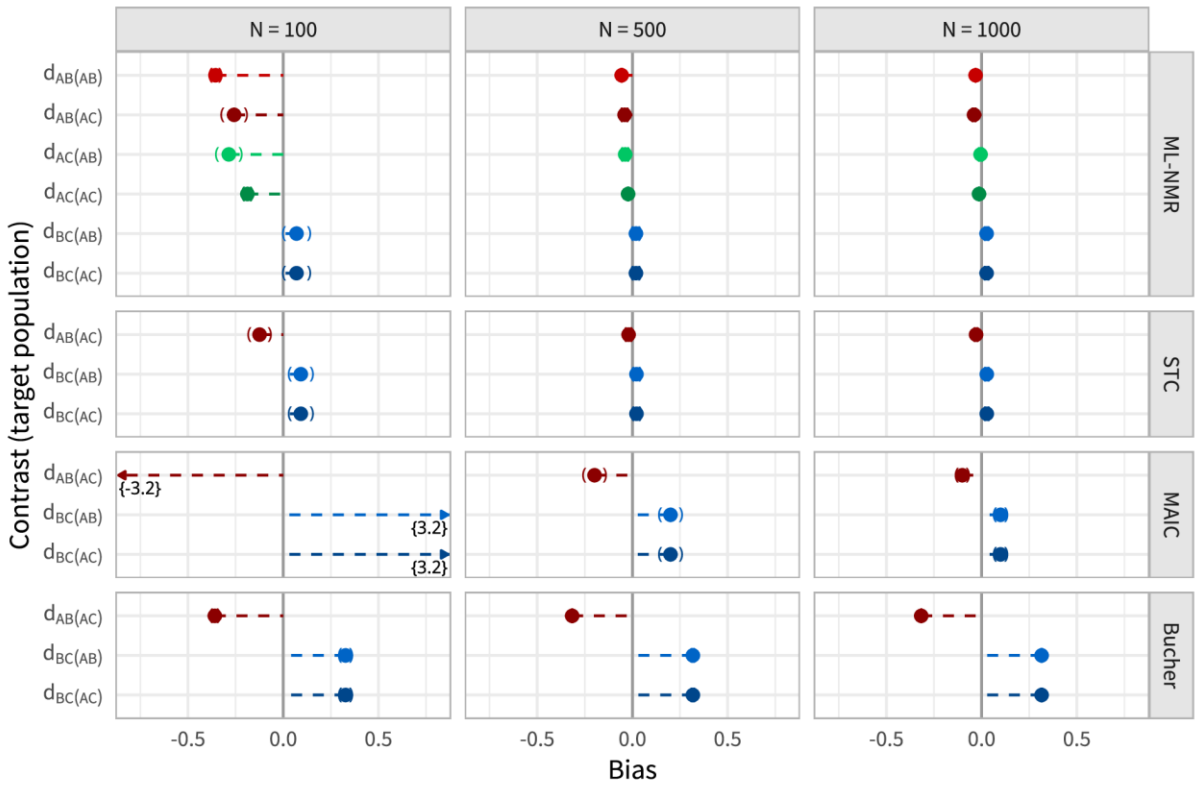
Aim: Assess the performance of population adjustment methods in a range of realistic scenarios under various failures of assumptions

Scenarios:

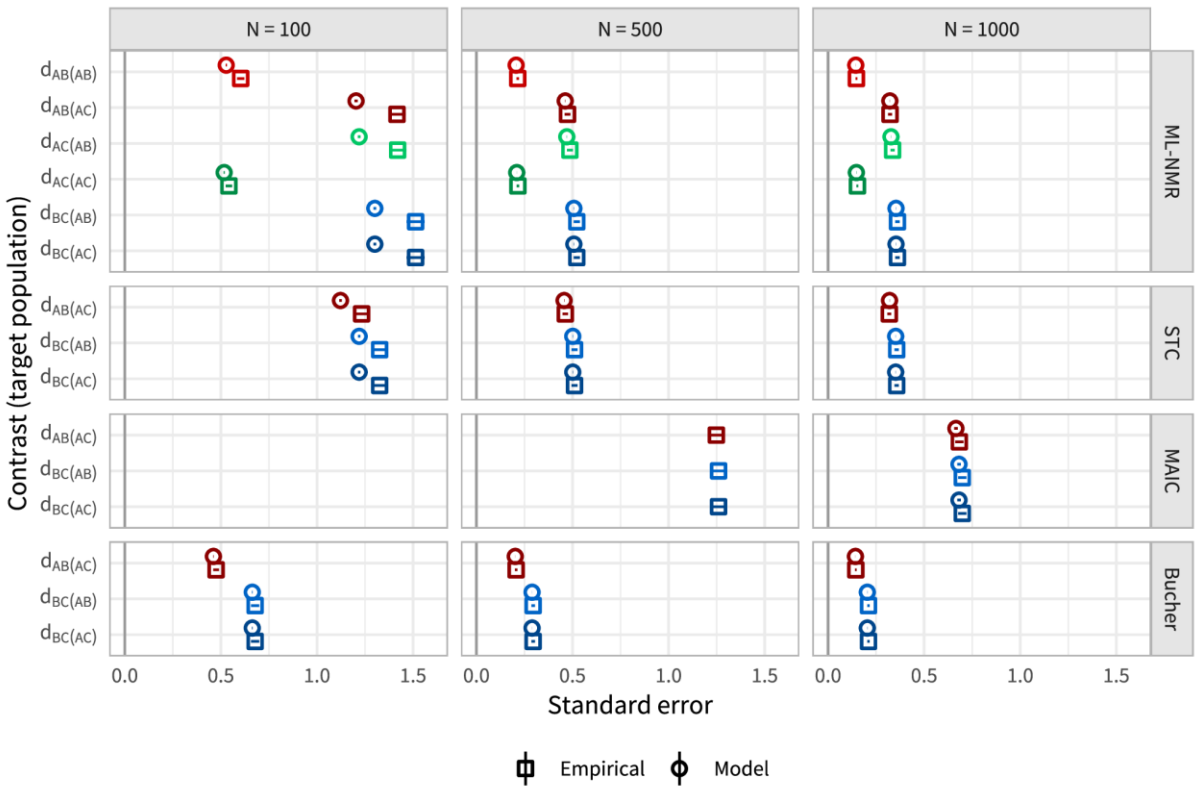
a.	Sample size
b.	Strength of effect modification
c.	Validity of shared EM assumption (common interactions)
d.	Strength of covariate correlations
e and f.	Validity of extrapolation/overlap
g, h, and i.	Validity of imputing joint covariate structure

Scenario a – sample size (all EMs)

Bias



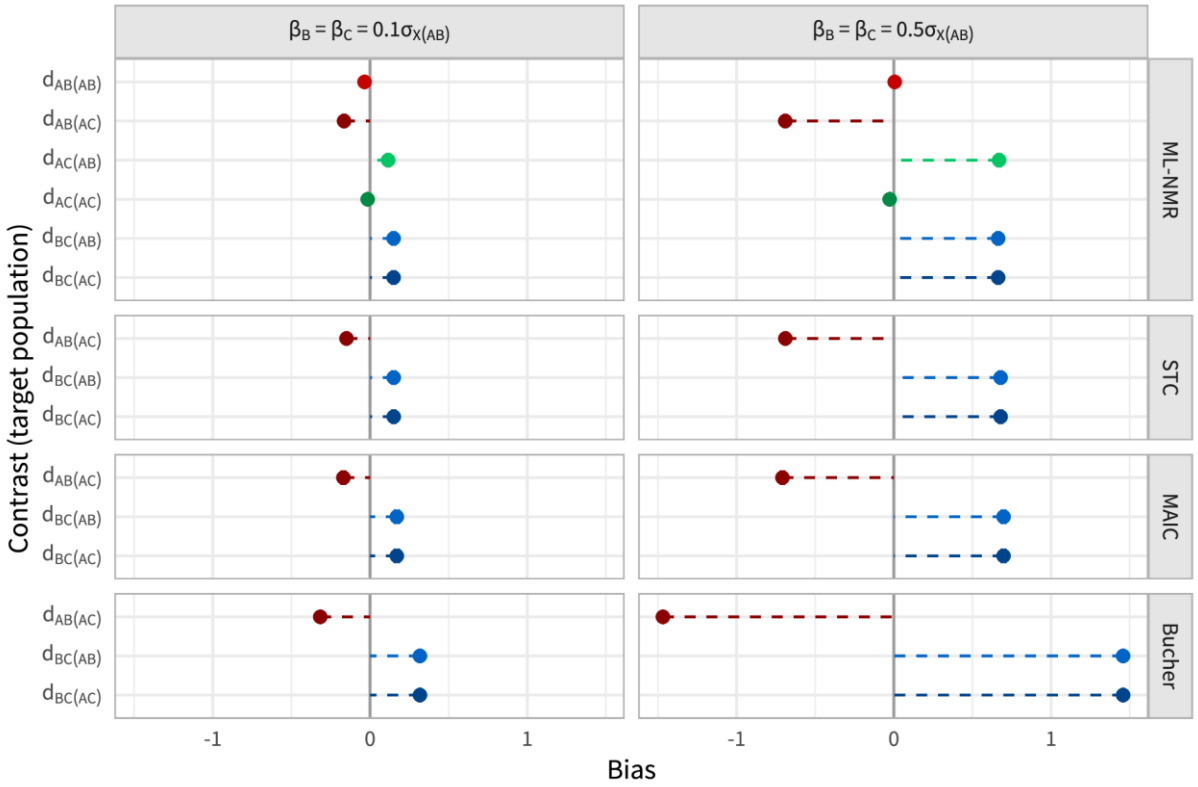
Standard Error



- ML-NMR and STC perform similarly at eliminating bias, some small sample bias (low average number of events)
- MAIC remains biased, even increases bias compared to standard Bucher IC with small N

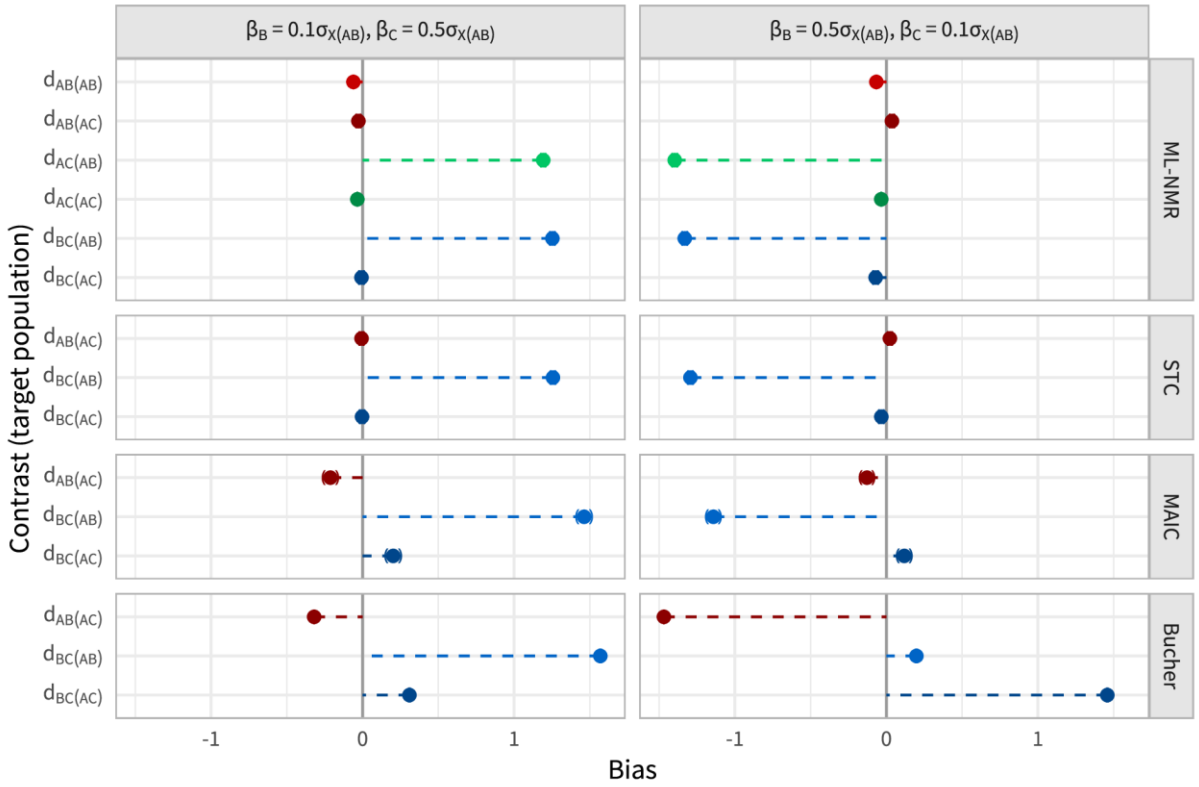
- ML-NMR and STC have similar standard errors, well-estimated, some underestimation with small N
- MAIC has larger standard error, bootstrap SE very unstable except at largest N

Scenario b – strength of effect modification
Bias (missing EM)



- All methods are biased when EMs are omitted
- The stronger the missing EM, the greater the bias
- Population adjustment methods otherwise unaffected by strength of EMs (not shown)

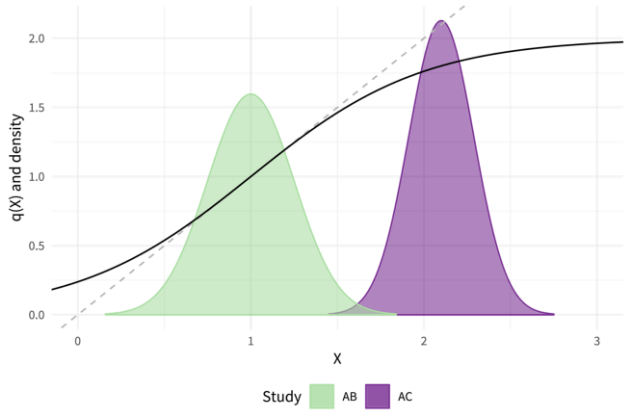
Scenario c – shared EM assumption
Bias (all EMs)



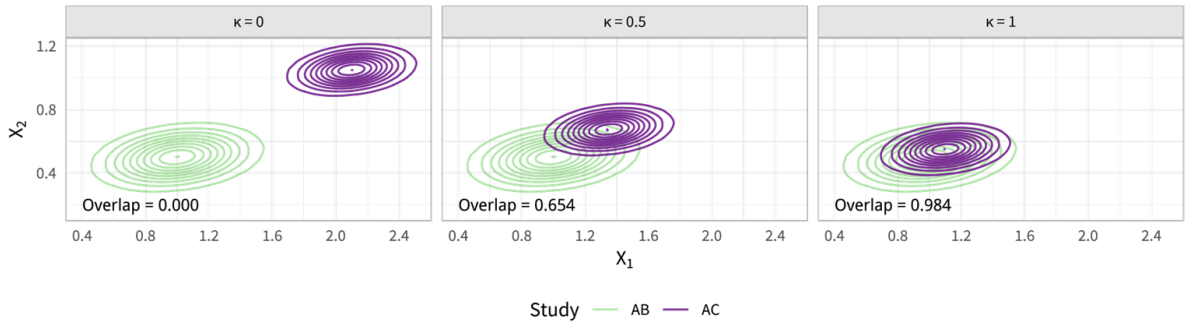
- Estimates of interaction term driven by IPD AB trial
- Estimates in AC population do not incur additional bias
- Estimates in AB (or any other) population incur bias

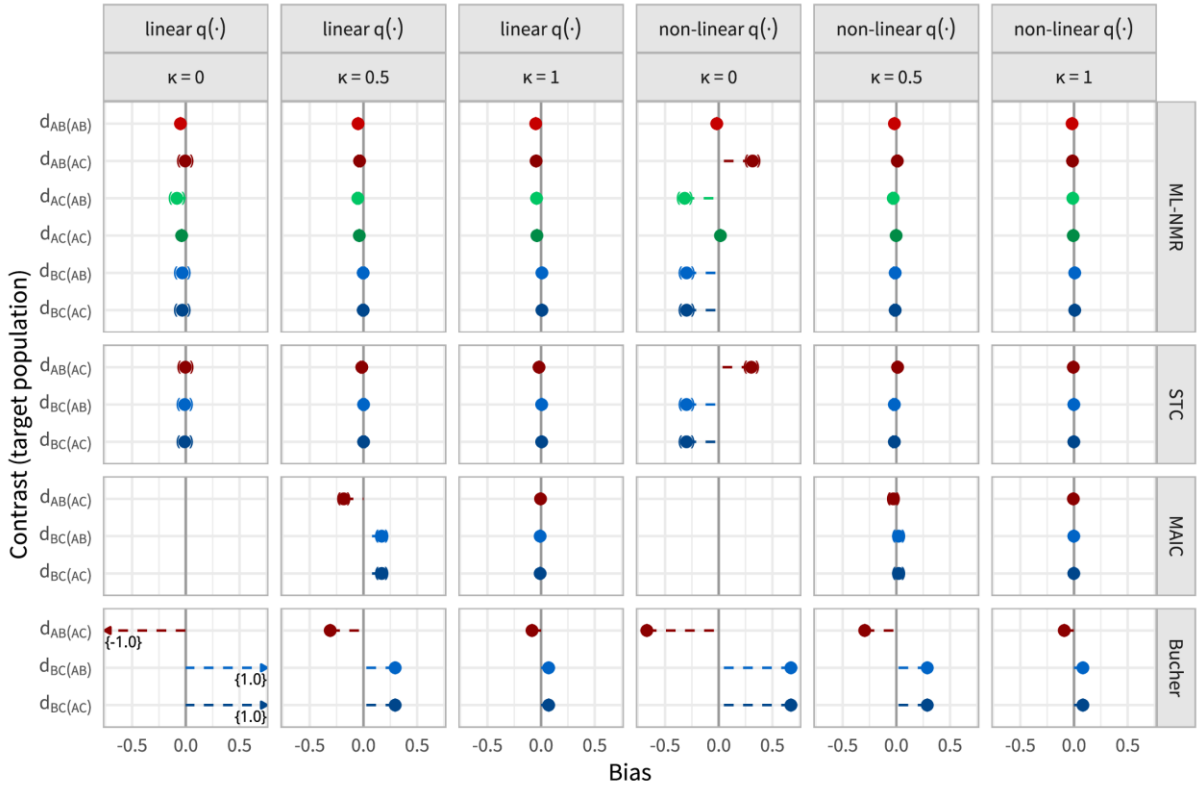
Scenarios e and f – validity of extrapolation and overlap

Linear or non-linear covariate-outcome relationship



Varying between-study overlap





- ML-NMR and STC unbiased when extrapolation is valid, biased when invalid (though still less than Standard IC here)
- MAIC cannot produce estimates when there is no overlap, estimates are only unbiased with full overlap, but then there is little bias to remove
- SEs for ML-NMR and STC approach Bucher as overlap increases, MAIC SEs only stable for full overlap (not shown)

University of BRISTOL

Conclusions

- ML-NMR and STC both performed very similarly throughout
  - Both incur bias when extrapolation or shared EM assumption invalid
  - Performance unchanged by covariate correlations, not sensitive to assumptions regarding joint covariate distribution in AgD study (scenarios d, g, h, and i, results not shown)
- MAIC performed poorly in almost all scenarios, in some cases even increasing bias compared to a standard indirect comparison
  - Especially with small sample sizes. Needs full overlap to be unbiased, and for stable estimation of SE
- All methods susceptible to bias (and resulting under-coverage) when missing any EMs
  - Highlights the need for careful, justified variable selection (TSD 18)
- ML-NMR offers additional advantages over previous methods
  - Combine IPD and AgD in networks of any size, relax/validate assumptions in larger networks
  - Numerical integration allows for general implementation regardless of model form or complexity
  - Decision making is aided by the production of estimates relevant to the decision target population

Further information

This work was undertaken with the support of the MRC grants MR/P015298/1 and MR/M005232/1, and the ConDuCT-II Hub (MR/K025643/1).

R package `multinma` on CRAN, doi: 10.5281/zenodo.3904454

Phillippo DM (2019) *Calibration of treatment effects in network meta-analysis using individual patient data.* PhD Thesis, University of Bristol. Available from research-information.bristol.ac.uk

Phillippo DM et al. (2020) *Multilevel Network Meta-Regression for population-adjusted treatment comparisons*. Journal of the Royal Statistical Society: Series A. doi: 10.1111/rssa.12579

Phillippo DM et al. (Under Revision) *Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study*

Phillippo DM et al. (2016) *NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparison in submissions to NICE*. NICE Decision Support Unit. Available from www.nicedsu.org.uk.